DOCUMENT RESUME

ED 110 170 PS 007 944

AUTHOR Kaufman, Maurice
TITLE The Effect of the DISTAR Instructional System: An Evaluation of the 1973-1974 Title I Program of Winthrop, Massachusetts.
PUB DATE [74]
NOTE 45p.; For the 1971-1972 report, see ED 070 525; for the 1972-1973 report, see PS 007 945

EDRS PRICE MF-$0.76 HC-$1.95 PLUS POSTAGE
DESCRIPTORS Achievement Gains; *Arithmetic; Comparative Analysis; Educationally Disadvantaged; *Instructional Programs; Language Development; Learning Disabilities; Mathematical Concepts; Oral Expression; *Primary Education; *Program Evaluation; *Reading Achievement; Self Concept; Student Attitudes; Test Interpretation; Word Study Skills
IDENTIFIERS *Distar; Elementary and Secondary Education Act Title I; ESEA Title I; Massachusetts

ABSTRACT

This report describes the evaluation of the 1973-1974 Distar I and Distar II programs, which were used as Title I programs for first and second graders. Distar I was used with two first grade classes; Distar II was used with two second grade classes. Children who had completed two years of Distar instruction received traditional instruction in third grade. The possible advantages (achievement in arithmetic computation) and disadvantages (inadequate word-meaning, paragraph-meaning, and arithmetic concept skills development), which were indicated in past comparative analyses of the Distar I and II programs, were reexamined. Findings suggest that: (1) the Distar program for first graders had no advantage over conventional instruction in developing oral language ability; (2) the second grade Distar program was strong in spelling, word study skills, and possibly in arithmetic computation; and (3) at the end of the third grade, Distar pupils appeared to perform approximately at grade level with distinct strength in arithmetic computation. Recommendations, including proposed steps toward redefining Title I populations and programs, are suggested. Appendixes contain data tables and other supportive material. (BRT)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED110170

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRO
DUCED EXACTLY AS RECEIVED FROM
THE PERSON OR ORGANIZATION ORIGIN
ATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRE
SENT OFFICIAL NATIONAL INSTITUTE OF
EDUCATION POSITION OR POLICY

The Effect of the DISTAR Instructional System: An Evaluation of the 1973-1974 Title I Program of Winthrop, Massachusetts

PS 007944

Maurice Kaufman, Ph.D.
Northeastern University

Boston

1974

00002

ABSTRACT

Distar I was used with two first-grade classes. Distar II was used with two second-grade classes. Control groups were selected at first- and second-grades to compare with Distar groups. Children who had completed two years of Distar instruction received traditional instruction in third-grade. These children were compared to control groups.

There was growth in oral language under both Distar and conventional first-grade instruction. First-grade reading results were inconclusive; the control group performed better than one Distar class but poorer than another. These children were retested in October 1974 under standardized test procedures. No significant differences among the classes were then obtained.

One Distar second-grade performed better than a control group in spelling and equally well in six other areas. The other Distar second-grade performed poorer in several areas than either group partially due to transfers of pupils, judged to be better achievers, from this group.

One Distar third-grade performed as well as the control group in four areas tested. The control group surpassed a second Distar third-grade in three areas. Pupils from the latter Distar group were matched by sex, age, and I.Q. with third-graders who attended the school in 1971. These pupils, who received conventional instruction, surpassed the Distar pupils in five of six areas.

Pupils in Distar and conventional programs have an equally favorable attitude toward school.

These results are compared to results attained in previous years, and general trends are identified. Recommendations, including proposed steps toward redefining Title I populations and programs, are made.

96693

CONTENTS PAGE

Introduction 1.

Description of the 1973-1974 Program 3.

Evaluation Design 4.

Tests Used 4.

Results 5.

- Oral Language 5.
- First-Grade Reading 6.
- Second-Grade Achievement 10.
- Third-Grade Achievement 14.
- Matched Third-Graders at Chase School 16.
- Pupil Attitude 19.

Summary of Results 21.

Conclusions 23.

Recommendations 24.

APPENDIX A: Interim Evaluation 26.

First-Grade Testing 27.

Second-Grade Testing 28.

Self-Concept Measure 29.

Parent Questionnaire 29.

Conclusions to Interim Report 31.

Tables to Interim Report 32-35.

APPENDIX B: October 1974 Retesting of 1st Graders 36.

Tables to APPENDIX B 38-41.

Introduction

For the past three years the Title I program at Winthrop consisted of the use of the Distar program in several first and second grade classes at E. B. Newton school (formerly at Center school) and at Chase school (formerly called Shirley Street school). The evaluation reports for 1971-1972 and 1972-1973 include descriptions of the published Distar program, reports of research employing the Distar program, and results of the program's use as the Winthrop Title I program.

In 1971-1972 the Title I program appeared to produce some gain in oral language ability of one first-grade Distar class. There appeared to be no advantage to Distar instruction over conventional instruction in first-grade reading achievement. A second-grade subgroup of Distar pupils performed significantly better in arithmetic computation than did a similar group given conventional instruction. All second-graders who had had Distar instruction were compared to national norms on the Stanford Achievement Test and were found to be, on the average, significantly below grade placement on four of the seven subtests.

The 1972-1973 report concluded that the Distar program conducted in that year appeared to have no advantage over conventional instruction in developing oral language ability of first-graders. Comparison of a Distar and non-Distar first-grade class at Newton school in reading achievement showed no significant differences. The first-grade Distar class at Shirley Street school appeared to perform better in reading than the other two classes. (No differences among these groups existed at the time of testing approximately six months later. See Appendix and Table A-4). Second-grade Distar pupils for the second year scored below grade level on Stanford Achievement subtests

of word meaning, paragraph meaning, and arithmetic concepts. At the end of third grade former Distar pupils did not perform as well in reading and arithmetic as their peers who received traditional instruction. The latter probably had higher initial ability. However, these third-grade former Distar pupils now scored above grade placement in arithmetic computation and were not significantly below grade placement on any Stanford Achievement subtests. Apparently these children overcame the significant discrepancy between grade placement and achievement that existed at the end of second grade on four subtests of the Stanford and were performing essentially at grade level in these areas after a year of conventional instruction.

Judging from the results obtained in 1971-1972 and 1972-1973 only, the use of the Distar program as Winthrop's Title I program appeared to have certain advantages (achievement in arithmetic computation) and disadvantages (perhaps temporary, as reflected in second-grade performance in word-meaning, paragraph meaning, and arithmetic concepts). These tentative judgments are reexamined in the present evaluation for 1973-1974. In addition, pupil attitude toward school is considered.

Discussions with Winthrop teachers and administrators suggest some differences of opinion concerning the intended population of the Distar program. Use of Distar as a Title I program can be supported in part because it is intended for children who are disadvantaged in a socio-economic sense. Title I requires that within target areas defined according to socio-economic criteria children who are educationally handicapped receive service funded by Title I. There is disagreement among the school staff as to whether Distar is appropriate for learning disabled children. The definition and identification of learning disabled children is itself a problem. Apart from

that, if Distar should be inappropriate for learning-disabled children, it becomes necessary to identify children who are educationally handicapped without being learning disabled and provide only these children with the Distar Title I program. On the other hand, if the population of educationally handicapped children includes or is synonymous with the population of learning disabled children, and the inappropriateness of Distar for learning disabled children is demonstrated, it is necessary to replace the Distar program with a more appropriate program for learning disabled children. The evaluation report will attempt to consider these two points: 1) is Distar appropriate or inappropriate for so-called learning disabled children? 2) does the population of educationally handicapped children include those who are learning disabled?

Description of the 1973-1974 Program

Distar I was used with one first-grade class at Newton school and one first-grade class at Chase school. For purposes of comparison, a control first-grade was selected at Dalrymple school (formerly called Highland Street school).

Distar II was used with one second-grade class at Newton school and one at Chase. Children in these classes had received instruction in Distar I as first graders. A control group at Newton school set up in 1972-1973 was continued as a control in 1973-1974.

Children who completed Distar II in 1972-1973 entered conventional third-grade classes in 1973-1974. Progress of these children was followed.

It was noted that Distar I instruction at Newton was supplemented by individual tutoring of various kinds given to about one-third of the group. Tutoring was given by remedial reading, learning disabilities, or speech specialists. At Chase, Distar I was supplemented within the classroom by materials from the Scott-

Foresman program. Distar II in both schools was typically supplemented by more conventional materials and instructional procedures.

Evaluation Design

The two first-grade Distar classes and the first-grade control class at Dalrymple were tested and compared on initial tests of readiness, oral language, I.Q., attitude, and on posttests of oral language and reading. Pretest scores of each class are compared to their posttest scores for evidence of growth in oral language and change in attitude.

The two second-grade Distar classes and the second-grade control class at Newton school received mid-year tests of reading and I.Q. Pre- and posttests of attitude were given. An end-of-year standardized achievement test of arithmetic, reading, and language was given. Between-group comparisons were made and pre-post comparisons of attitude were made for each group.

At mid-year, parents of first- and second-graders were asked to respond to a questionnaire concerning their impression of the instructional program. Comparisons of responses by parents of Distar and non-Distar pupils are made in the Appendix.

Third-grade former Distar and former control pupils are compared on end-of-year performance on a standardized achievement test of arithmetic and reading. Third-grade former Distar pupils at Chase were matched with children who were in third-grade in 1971 on the basis of sex, I.Q., and chronological age at time of testing. Performance on a standardized achievement test by each group is compared.

Tests Used

First-grade oral language was measured by means of the First-Grade Oral Language Test devised by the evaluator. Development of this test is described in the 1972-1973 report (p. 6).

An evaluator-designed reading test was used for midyear second-grade testing and end-of-year first-grade testing. This is a shortened version of the First-Grade Reading Test used in prior evaluations. It consists of six selections, six questions on main idea, twelve on stated details, and six on inferences for a total of twenty-four items.

Readiness testing was done with the Metropolitan Readiness Test. The Otis-Lennon Mental Ability Test was used for IQ measurement. The Stanford Achievement Test, Primary II, was used for end-of-year testing of second and third-graders. The attitude measure was When Do I Smile? Lower Elementary Level, published by American Institutes for Research.

Results

Oral Language

Table 1 shows the first-grade Distar class at Newton school and the first-grade control class at Dalrymple made significant gains in oral language ability as measured on the Oral Language Test. The Distar class at Newton and the control class both performed significantly better on the posttest than the first-grade class at Chase. This is shown in Table 2. A comparison of these three groups on the pretest is made in Table A-3. Differences in number of children tested in any one school as shown on the three tables result from the fact that a child may have been present for one test but not for another. Table 1 includes only those children both pre- and posttested. This variation requires that any comparison of results from one table to another be made with caution.

TABLE 1

PRETEST AND POSTTEST RESULTS ON THE FIRST-GRADE ORAL LANGUAGE TEST

Group	N	Pretest Mean & SD	Posttest Mean & SD	t	P
Distar-Newton	21	7.86 2.61	10.62 2.50	3.610	<.01
Distar-Chase	18	8.44 2.23	7.61 3.36	-1.056	NS
Control-Dalrym.	20	10.10 2.61	11.50 2.37	3.339	<.01

NS = Not Significant

TABLE 2

COMPARISON OF ORAL LANGUAGE POSTTEST SCORES OF FIRST-GRADERS

Group	N	Mean	SD	t	P
Newton-Distar	22	10.73	2.49	3.573	<.001
vs.					
Chase-Distar	20	7.55	3.25		
Newton-Distar	22	10.73	2.49	-0.857	NS
vs.					
Dalrym.-Contr.	23	11.35	2.37		
Chase-Distar	20	7.55	3.25	-4.418	<.001
vs.					
Dalrym.Contr.	23	11.35	2.37		

NS = Not Significant

First-Grade Reading

Although many significant differences appear in a comparison of the three first-grade classes with one-another in reading ability, from the standpoint of judging the effectiveness of the Distar program, results are inconclusive and ambiguous. The reason for this is the marked disparity in performance of the two Distar classes. As shown in Table 3, performance of the Distar class at Chase was

decidedly superior to that at Newton. Not only are the means of the Chase group higher but the spread of scores, as measured by the standard deviation, is consistently smaller. This fact is also illustrated in Table 4. The same results were obtained when comparing Chase with the control group. There are several possible explanations for this result:

Eleven of the twenty-one children at Newton are receiving supplementary remedial reading or learning disability help, or have been identified for such help in the future. It is possible that this is disproportionately large, compared to the other two classes, making for a basically dissimilar group with lower achievement potential.

A large number of the test papers at Newton were incorrectly marked. There were numerous omissions, multiply-marked answers, or marks in wrong places (other than choices in the multiple-choice questions). Similarly, these children failed to mark the Smile pretest correctly, resulting in the invalidation of that instrument for many children. These observations suggest either that the children were very immature and unable to learn the mechanics of responding to paper and pencil instruments, or that they were poorly supervised during the testing sessions.

The instructional program conducted at Chase differed from that at Newton by combining many aspects of the Scott-Foresman program with the Distar program. The results that were obtained may be due to this fact, in which case such a combination program must be superior to Distar alone and Scott-Foresman as practiced in the control group. On the other hand, it is necessary to recall the results of the program in 1972-1973. The first-grade at Chase apparently did extremely well at that time. With few changes, these children as second-graders are compared to second-graders at Newton on the New First-Grade Reading Test administered in

December 1973. These results appear in Table A-4. It can be seen that despite apparent differences at the end of first-grade, there are no significant differences favoring the Chase pupils six months later. If the apparently superior achievement of the first-graders at Chase is, likewise, not retained in six month's time, the educational importance of this difference must be questioned.

Finally, the possibility that children at Chase were inadvertently coached as teachers supervised the mechanics of test administration cannot be ruled out. Inadvertent coaching at Chase along with poor supervision at Newton could explain the large differences in test performance.

Several recommendations are in order: 1) Assuming the superior test performance at Chase reflects superior reading ability, the second-grade teacher of these children should be encouraged to support their continuous growth. 2) If there were differences in the kinds of test supervision among the three classes, future group tests of first-graders should be administered by the same examiner team, perhaps consisting of a remedial reading teacher or reading supervisor trained to administer group tests assisted by aides. 3) To check on maintenance of performance level, the same group test should be administered to these children when they are in second-grade in the fall of 1974.

TABLE 3

COMPARISON OF READING SCORES ON
NEW DISTAR READING TEST
FOR FIRST GRADERS

Score	Newton-Distar (N = 21) Mean & SD	Chase-Distar (N = 19) Mean & SD	t	P
Main Idea	1.24 1.73	4.32 1.11	-6.621	<.001
Details	1.81 2.52	10.16 1.42	-12.701	<.001
Inferences	0.86 1.35	3.68 1.25	-6.844	<.001
Total Reading	3.90 4.49	18.16 2.41	-12.308	<.001
	Newton-Distar (N = 21) Mean & SD	Dalrym.-Contr. (N = 23) Mean & SD		
Main Idea	1.24 1.73	2.91 1.56	-3.374	<.01
Details	1.81 2.52	6.52 3.01	-5.596	<.001
Inferences	0.86 1.35	3.13 1.42	-5.418	<.001
Total Reading	3.90 4.49	12.57 4.97	-6.041	<.001
	Chase-Distar (N = 19) Mean & SD	Dalrym.-Contr. (N = 23) Mean & SD		
Main Idea	4.32 1.11	2.91 1.56	3.284	<.01
Details	10.16 1.42	6.52 3.01	4.826	<.001
Inferences	3.68 1.25	3.13 1.42	1.325	NS
Total Reading	18.16 2.41	12.57 4.97	4.481	<.001

NS = Not Significant

TABLE 4

FREQUENCY DISTRIBUTION OF SCORES ON FIRST-GRADE READING TEST

Subtest	Scores	Newton f	Chase f	Dalrym. f
Main Idea	5 - 6	2	9	5
	3 - 4	3	10	8
	0 - 2	16	0	11
Details	9 - 12	1	18	7
	5 - 8	1	1	10
	0 - 4	19	0	7
Inferences	5 - 6	0	5	2
	3 - 4	3	9	15
	0 - 2	18	5	7
Total Reading	17 - 24	1	13	6
	9 - 16	1	6	13
	0 - 8	19	0	5

Second-Grade Achievement

The interim report (Appendix) shows that the three second-grade classes were essentially equivalent in reading ability in December as measured on the new form of the evaluator-constructed Distar Reading Test. Only one difference was found between the Newton Distar class and the Chase Distar class favoring the former. Of a possible score of 24, Table A-4 shows the Newton Distar group attained a total mean score of 19.87, the Chase Distar group attained a total mean score of 18.16, and the Newton Control group a total mean score of 18.23.

These groups were again tested in May 1974. Membership in the groups was not identical from December to May in that a few children were tested on one occasion but not on the other. It should be noted that three children included at Chase in the December testing had been transferred to traditional classes at the beginning of the school year.

These children are not included in the May testing.

Table 5 compares the three groups on the Stanford Achievement Test, Primary II, subtests of Word Meaning, Paragraph Meaning, Spelling, Word Study Skills, Language, Arithmetic Computation, and Arithmetic Concepts. Subtests of Spelling and Language were not administered at Chase school; no comparisons with Chase can be made for these areas. Subtest comparisons are presented in Table 5.

As was stated in the 1972-1973 report, the control group was selected in such a manner as to permit comparisons with the Distar group at Newton school. Furthermore, additional changes in the Chase school group, including movement of three students who achieved well into traditional programs, make inappropriate any evaluative judgments involving Chase school in the comparison made on Table 5. The data permit one to assess only the present status of the Chase school group without inferring anything about the effectiveness of the instructional program. On this basis it can be seen that the Chase school group at the end of the second grade consisted of children who performed poorer in word meaning, paragraph meaning, and word study skills than children in each of the other two groups.

Comparison of Distar and Control groups at Newton school shows the former performed significantly better on the spelling subtest. There were no other significant differences between the two groups. It can be concluded that the Distar program at Newton school was at least as effective as the conventional program for these children. The Distar program evidently produced superior results in spelling.

Grade equivalent scores on the Stanford subtests were compared to a grade placement norm of 2.9. As shown in Table 6, the Newton school Distar class performed significantly better than the national norms in spelling and word study skills; the Newton school control class scored significantly below national norms in language; the Chase school Distar class scored significantly below national norms

in word meaning, paragraph meaning, and word study skills.

TABLE 5

COMPARISON OF DISTAR AND NON-DISTAR PUPILS
ON STANFORD ACHIEVEMENT TEST
PRIMARY II, FORM W (MAY 1974)

Test	Newton Distar Mean & SD	N	Chase Distar Mean & SD	N	t	P
Word Meaning	2.93 .67	23	2.29 .50	15	3.163	<.01
Para. Meaning	2.80 .77	23	2.19 .47	15	2.755	<.01
Word St. Skills	3.70 1.52	23	2.49 .69	14	2.786	<.01
Arith. Compu.	3.03 .51	23	3.00 .21	13	0.175	NS
Arith. Concept	2.92 1.03	23	2.67 .53	13	0.822	NS

	Newton Distar Mean & SD	N	Newton Control Mean & SD	N		
Word Meaning	2.93 .67	23	2.94 .55	27	-0.035	NS
Para. Meaning	2.80 .77	23	2.89 .57	27	-0.487	NS
Spelling	3.31 .62	23	2.73 .57	27	3.471	<.01
Word St. Skills	3.70 1.52	23	3.49 1.58	27	0.470	NS
Language	2.97 .84	23	2.63 .64	27	1.579	NS
Arith. Compu.	3.03 .51	23	2.74 .53	27	1.929	NS
Arith. Concept	2.92 1.03	23	2.74 .87	27	0.687	NS

Test	Chase Distar Mean & SD	N	Newton Control Mean & SD	N	t	P
Word Meaning	2.29 .50	15	2.94 .55	27	-3.780	<.001
Para. Meaning	2.19 .47	15	2.89 .57	27	-4.096	<.001
Word St. Skills	2.49 .69	14	3.49 1.58	27	-2.244	<.05
Arith. Compu.	3.00 .21	13	2.74 .53	27	1.692	NS
Arith. Concept	2.67 .53	13	2.74 .87	27	-0.258	NS

NS = Not Significant

TABLE 6

COMPARISON OF STANFORD ACHIEVEMENT SUBTEST SCORES
TO GRADE PLACEMENT NORM OF 2.9

Test	Newton Distar Mean & SD	N	P	Chase Distar Mean & SD	N	P	Newton Contr. Mean & SD	N	P
Word Meaning	2.93 .67	23	NS	2.29 .50	15	<.001	2.94 .55	27	NS
Para. Meaning	2.80 .77	23	NS	2.19 .47	15	<.001	2.89 .57	27	NS
Spelling	3.31 .62	23	<.01				2.73 .57	27	NS
Word St. Skills	3.70 1.52	23	<.05	2.49 .69	14	<.05	3.49 1.58	27	NS
Language	2.97 .84	23	NS				2.63 .64	27	<.05
Arith. Compu.	3.03 .51	23	NS	3.00 .21	13	NS	2.74 .53	27	NS
Arith. Concept	2.92 1.03	23	NS	2.67 .53	13	NS	2.74 .87	27	NS

NS = Not Significant. This comparison employed t test for single mean.

Third-Grade Achievement

Third-grade former Distar pupils from Newton and Chase schools and former control pupils at Dalrymple (Highland Street) were compared on results of system-wide testing on the Stanford Achievement Primary II Form W. These pupils were among those in the second-grade group evaluated in 1972-1973. Distar pupils were not combined as one group, as done in 1972-1973, to reveal possible differences between the two former Distar classes at the present time.

All children included in this analysis were tested on the word meaning, paragraph meaning, arithmetic computation, and arithmetic concepts subtests of the Stanford. Results of the comparison are presented in Table 7. Table 7 reveals that several significant differences appear between the two former Distar groups, favoring the group from Newton school. No differences were obtained between the Newton Distar group and the Dalrymple controls, but the comparison of the Dalrymple controls with the Chase school former Distar pupils revealed significant differences favoring the former in word meaning and paragraph meaning.

These results were not inconsistent with the second-grade comparison of 1972-1973 in which essentially the same children were compared. In that comparison the control group was favored in word meaning and word study skills. Evidently, the particular Distar school has some influence on performance. Newton school attained better results in the present year's third grade and second grade than children at Chase. This may reflect differences in first-grade selection policy, in policy for transferring children from Distar to traditional programs in second-grade on the basis of superior first-grade performance, and in quality of instruction in Distar classes in grades one and two and in post-Distar classes in grade three.

TABLE 7

COMPARISON OF THIRD GRADE FORMER DISTAR & CONTROL PUPILS ON STANFORD ACHIEVEMENT TEST, PRIMARY II, FORM W (APRIL 1974)

Test	Newton Distar Mean & SD	N	Chase Distar Mean & SD	N	t	P
Word Meaning	3.78 .59	18	3.16 .81	23	2.732	<.01
Para. Meaning	3.71 .73	18	2.89 .93	23	3.052	<.01
Arith. Compu.	4.41 .81	18	3.80 1.22	23	1.835	NS
Arith. Concept	4.04 .75	18	3.25 1.02	23	2.758	<.01
	Newton Distar Mean & SD	N	Dalrym. Contr. Mean & SD	N		
Word Meaning	3.78 .59	18	3.92 .82	13	-0.552	NS
Para. Meaning	3.71 .73	18	3.53 .55	13	0.728	NS
Arith. Compu.	4.41 .81	18	4.10 .71	13	1.113	NS
Arith. Concept	4.04 .75	18	3.75 1.18	13	0.837	NS
	Chase Distar Mean & SD	N	Dalrym. Contr. Mean & SD	N		
Word Meaning	3.16 .81	23	3.92 .82	13	-2.692	<.05
Para. Meaning	2.89 .93	23	3.53 .55	13	-2.259	<.05
Arith. Compu.	3.80 1.22	23	4.10 .71	13	-0.811	NS
Arith. Concept	3.25 1.02	23	3.75 1.18	13	-1.335	NS

Mean subtest scores of third graders were compared to grade expectancy norm of 3.7. This comparison (Table 8) shows that the Newton school Distar group significantly surpassed the norm of 3.7 in arithmetic computation. The Chase school Distar group scored significantly below the expected score of 3.7 in word meaning, paragraph meaning, and arithmetic concepts. The single area of success for both groups, then, was arithmetic computation. The Dalrymple school control group scored approximately at the national norm of 3.7 in all areas tested.

TABLE 8

COMPARISON OF STANFORD ACHIEVEMENT SUBTEST SCORES TO GRADE PLACEMENT NORM OF 3.7

Test	Newton Distar N = 18 Mean & SD	P	Chase Distar N = 23 Mean & SD	P	Dalrym. Contr. N = 13 Mean & SD	P
Word Meaning	3.78 .59	NS	3.16 .81	<.01	3.92 .82	NS
Para. Meaning	3.71 .73	NS	2.89 .93	<.001	3.53 .55	NS
Arith. Compu.	4.41 .81	<.01	3.80 1.22	NS	4.10 .71	NS
Arith. Concept	4.04 .73	NS	3.25 1.02	<.05	3.75 1.18	NS

Matched Third Graders at Chase School

It was the impression of the principal at Chase school that the Distar program was more successful than the conventional program when used with children thought to have potential learning problems. This expectation was evaluated in two ways. First, a cursory examination was made of first and second grade results on the Gates-MacGinitie reading tests, which are used regularly at that school. The first-grade Distar class obtained a mean score of 1.9 in vocabulary and 1.6 in comprehension. These were lower than mean scores obtained by the

other two first-grade classes. The second grade Distar class obtained mean scores of 2.6 in vocabulary and 2.0 in comprehension. These, too, were lower than the mean scores of the other two second-grade classes. These results indicate the children selected for Distar instruction at Chase school do not perform as well as other children at Chase school. This may be due to the selection process for Distar; children who are judged to have learning problems are perhaps more often selected. Also, children who achieve well in first-grade are often not retained in Distar in second-grade, thereby reducing the mean achievement of that group. In any event, the Distar program has not enabled children selected for the program to achieve as well as other children in the grade. The major criterion for judging the success of the effort, however, is whether children selected for the program perform better in the program than they would perform in the alternative, conventional program. The second step was taken in an effort to answer this question.

An attempt was made to match the twenty-three former Distar third-graders (whose results on four Stanford subtests are reported in the preceding section of this report) with third-graders in 1971 who had proceeded through a conventional instructional program. One child had not started Distar in September of the first-grade. Of the remaining twenty-two children, eighteen were successfully matched with children from three third-grade classes at Chase in 1971 on the basis of sex, I.Q. (second-grade administration of the Otis) within five I.Q. points, and chronological age (at time of I.Q. testing) within six months. The matched groups were then compared on six Stanford subtests using the t test for correlated observations. As shown on Table 9, the two groups did not differ in I.Q. and chronological age (at time of testing), therefore, they were successfully matched. The non-Distar group achieved better

than the Distar group on subtests of word meaning, paragraph meaning, arithmetic concepts, word study skills, and spelling. The groups did not differ significantly in arithmetic computation. This procedure of analysis was followed in order to judge how well Distar children might have performed under the conventional instructional program. It is possible that, despite the effort of matching, the control group was not comparable. The basis for selection of children, in the first place, for Distar instruction is not known. First-grade readiness scores of the two groups were not examined and may differ. The assumption of similarity between groups on the basis of second grade I.Q. was made. In any event, the evidence at this time does not support the likelihood that the third-grade children at Chase school who had been selected for Distar instruction achieved any better than they would have achieved in a conventional program of instruction.

TABLE 9

COMPARISON OF EIGHTEEN PAIRS OF DISTAR THIRD-GRADERS IN 1974 AND NON-DISTAR THIRD GRADERS IN 1971 AT CHASE SCHOOL

	Distar Mean & SD	Non-Distar Mean & SD	t	P
Age in mos. at time of IQ test	108.06 3.69	108.33 3.46	0.424	NS
I.Q.	101.17 10.39	100.89 9.68	-0.417	NS
Stanford Achievement Prim. II, W Subtest Scores				
Word. Mng.	3.32 .76	4.18 .59	3.806	<.01
Para. Mng.	3.08 .95	4.12 .61	4.441	<.001
Arith. Comp.	3.99 .98	4.12 .73	0.442	NS

TABLE 9 (CONTINUED)

	Distar Mean & SD	Non-Distar Mean & SD	t	P
Arith. Conc.	3.42 1.03	4.14 .90	2.370	<.05
Wd. St. Sk.	3.68 1.61	5.22 1.50	2.785	<.05
Spelling	3.48 .88	4.42 1.05	2.560	<.05

NS = Not Significant

Pupil Attitude

It was considered possible that participation in the Distar program would have some beneficial effect upon a pupil's attitude toward school, which reflects to some extent a pupil's self-concept. The instrument When Do I Smile? Lower Elementary Level (American Institutes for Research) was selected to measure attitude. Pretesting was done in November 1973. A comparison of first-grade classes and second-grade classes is presented in the Appendix (Table A-5). A comparison of the first-grade Distar class at Chase with the first-grade control class at Dalrymple showed the latter group had a more favorable attitude (lower mean score) on the pretest.

A comparison of posttest scores shows no significant differences between any two first-grades or between any two second-grades (see Table 10). Comparisons of pretest with posttest results for each group is made in Table 11. As shown there, a significant change toward a less favorable attitude occurred in the first-grade control group. However, as shown in Table 10, this group (of twenty-three children on the post-test compared to nineteen of the same children in the pre-post comparison) did not have a less favorable attitude than either of the Distar groups on the posttest. In conclusion, the Distar program did not produce either a more or a less favorable attitude toward school than a conventional instructional

program.

The Smile instrument has a possible raw score range of 21 to 105. A score of < 63 would result from expressions of favorable attitude. It can be seen that all mean scores reported in Tables 10 and 11 indicate favorable attitudes toward school.

TABLE 10

COMPARISON OF GROUPS ON ATTITUDE MEASURE
WHEN DO I SMILE? POSTTEST (MAY, 1974)

Group	N	Mean	SD	t	P
Grade 1					
Newton-Distar	17	37.41	12.02	-1.903	NS
vs. Chase-Distar	17	44.94	11.03		
Newton-Distar	17	37.41	12.02	-1.416	NS
vs. Dalrym.-Contr.	23	43.57	11.82		
Chase-Distar	17	44.94	11.03	0.374	NS
vs. Dalrym.-Contr.	23	43.57	11.82		
Grade 2					
Newton-Distar	22	37.55	8.81	-1.166	NS
vs. Chase-Distar	15	41.60	12.37		
Newton-Distar	22	37.55	8.81	-0.252	NS
vs. Newton-Contr.	25	38.12	6.81		
Chase-Distar	15	41.60	12.37	1.151	NS
vs. Newton-Contr.	25	38.12	6.81		

NS = Not Significant
Lower scores indicate more-favorable attitude

TABLE 11

PRETEST AND POSTTEST RESULTS ON THE ATTITUDE MEASURE WHEN DO I SMILE?

Group	N	Pretest Mean & SD	Posttest Mean & SD	t	P
Grade 1					
Newton-Distar	9	39.22 17.73	38.33 13.89	-0.139	NS
Chase-Distar	16	44.44 12.37	45.19 11.34	0.188	NS
Dalrym.-Contr.	19	32.84 10.08	42.74 11.63	2.903	<.01
Grade 2					
Newton-Distar	21	38.95 12.91	37.14 8.82	-0.648	NS
Chase-Distar	13	40.86 9.21	41.86 12.80	0.317	NS
Newton-Contr.	23	39.09 9.65	38.43 7.01	-0.306	NS

NS = Not Significant
Negative t value indicates shift to more favorable attitude

Summary of Results

1. The Distar first-grade at Newton school and the first-grade control at Dalrymple made significant gains in oral language. The two classes performed better on the posttest than the Distar first-grade at Chase.

2. The control first-grade at Dalrymple performed better than the Distar first-grade at Newton but poorer than the Distar first-grade at Chase. Several possible explanations for this result are discussed, including that of error in test administration.

3. The Newton-Distar second grade performed better than the Newton-Control second grade on the spelling subtest of the Stanford Achievement. There were no differences between these two groups on six other subtests. The Chase school Distar class, partly due

to interclass transfers, consisted of children who performed poorer in word meaning, paragraph meaning and word study skills than the other groups.

4. The Newton-Distar second-grade surpassed the grade placement norm of 2.9 in spelling and word study skills. The Newton-control second-grade class scored below grade placement in language. The Chase-Distar second-grade scored below the national norm in word meaning, paragraph meaning, and word study skills.

5. The Newton former Distar third-grade and Dalrymple control third-grade did not differ significantly on four Stanford Achievement subtests. The former surpassed the Chase former Distar third-grade in word meaning, paragraph meaning, and arithmetic concepts. The Dalrymple control surpassed Chase in word meaning and paragraph meaning.

6. The Newton former Distar third-grade surpassed the national norm of 3.7 in arithmetic computation. The Chase former Distar third-grade fell significantly below the norm in word meaning, paragraph meaning, and arithmetic concepts.

7. Eighteen third-grade former Distar pupils at Chase were matched with children who were third-graders in 1971 for sex, I.Q., and chronological age (when tested). The former Distar pupils scored significantly below the others in five of six Stanford Achievement subtests.

8. First- and second-grade Distar and control classes showed, on the average, favorable school attitude. There was a significant shift toward a less favorable attitude in the first-grade control class, but neither the first-grade classes nor the second-grade classes differed significantly on the post-tested attitude measure.

Conclusions

1. There was evidence of growth in oral language ability under both Distar and conventional first-grade instruction. Over the past three years there is growing evidence that the Distar program has no advantage over conventional instruction in developing oral language ability of first-graders.

2. For the second year, performance of first-graders at Chase surpassed other first-graders in reading. The first-graders in 1972-1973 did not maintain this advantage. It is not clear whether the difference is due to features of the Distar instructional program at Chase, error in test administration procedures, or some other reason. Unlike past years, the first-grade Distar class at Newton school performed poorer in reading than the control group. There was evidence that the children in the Newton group were insufficiently supervised on the reading test and on the attitude pretest. Also, a substantial number of the Newton school children had been identified as having learning disabilities. It appears likely, judging from test results and teacher observations, that a population so identified fails to achieve well in a Distar program and requires some alternative instructional program.

3. At second-grade level Distar pupils appear to do as well as their peers of similar initial ability in control groups. Second graders in Distar, when pupils of higher ability are retained in Distar classes, have begun to score at and above national norms for their grade placement. At second-grade level, distinct strengths of such a Distar group (as was found in Newton school) seem to lie in spelling, word study skills, and possibly in arithmetic computation.

4. Third-graders who had Distar instruction in the first one or two grades appear to differ depending on their school. Former Distar pupils at Newton evidently achieve better than those at Chase in several areas. Distar pupils at Newton in 1973-1974 perform as well as children who had proceeded through conventional programs of instruction. The Newton school former Distar group surpassed national norms in arithmetic computation and were at grade level in other areas. Except for the current year's third-grade at Chase, Distar pupils appear to perform approximately at grade level by the end of third-grade with a distinct strength in arithmetic computation. Pupils at Chase of comparable I.Q. appear to have performed better after three years of conventional instruction.

5. Pupils in Distar and conventional programs appear to have an equally favorable attitude toward school.

Recommendations

1. Largely to check on the ambiguous status of first-grade reading results, the same children should be retested on the same reading test in the fall of 1974 with safeguards to assure comparable procedures of test administration.

2. Equivalent criteria should be established for both target schools (Chase and Newton) to employ in selecting children for Title I service on the basis of educational handicap. This is especially important for pre-first-graders, where achievement data is lacking. Within this group of educationally handicapped children are those who have been referred to as learning disabled. Criteria for the definition and identification of this subgroup should be established. Since it is likely that all educationally handicapped children do not benefit from the Distar instructional program, the Winthrop schools should begin to use Title I funds to support a broader range of instructional services. Instruction should be differentiated

according to the characteristics of the learner. The form of instruction provided to a child should be continued only if progress day-by-day is actually observed. It is unnecessary to wait until end-of-year evaluation to make such a judgment.

3. To implement the recommendations in paragraph (2), the school department should a) conduct a needs assessment, b) establish criteria for the selection of children for Title I service on the basis of educational handicap, c) establish a range of instructional services related to the types of educational handicap identified within the population, d) establish a set of learning goals of a highly specific type so that day-to-day progress can be evaluated, e) identify, by name, the children who will receive Title I service annually, f) justify, on the basis of objective and generally accepted criteria, non-inclusion of a child who had received Title I service the previous year.

4. The Distar program should be used only with those children who are observed to progress under it on a regular basis. The program should be supplemented so that a wider range of reading and writing activities will be possible. There should be increased opportunity to develop comprehension through the reading of stories and factual articles, greater exposure to literature, and increased opportunity to write and read about experiences shared by the children.

5. All children should have the opportunity to use the materials of many-faceted basic reading systems that are now being introduced into the schools.

APPENDIX A

Interim Evaluation of the 1973-1974 Title I Program of Winthrop, Massachusetts

Maurice Kaufman, Ph.D.

Boston

1974

First Grade Testing

The testing of first-graders to date was designed for two purposes: 1) to describe and compare the Distar and Control groups with respect to IQ, readiness, and oral language ability; 2) to establish a baseline for future growth in oral language ability.

Examination of Table A-1 shows that the two first-grade Distar classes are approximately equal in IQ. The first-grade control class at Dalrymple School is superior to one of the Distar classes. This superiority may reflect a difference in native ability or it may reflect a deficiency in performance of the Distar children that is a direct result of the socio-economic deprivation whose effects on academic performance the Title I program proposes to overcome.

Similarly, the Dalrymple group was superior to the two Distar classes on the Metropolitan Readiness Test (Table A-2). The two Distar groups did not differ significantly in readiness.

The Oral Language Test was used for the purpose of comparing the groups in initial oral language ability, and to secure a base line to judge growth in oral language during the year. The test will be readministered in May at which time the groups will again be compared and, in addition, the growth within each group will be measured. The initial Oral Language Test results show (Table A-3) that the control group at Dalrymple School was again superior to either Distar class. The Distar classes performed equally well. It is to be hoped that, if instruction in oral language development is effective, the Distar group will do as well as the control group at the end of the year. In addition, the Distar groups should do as well as the control group on a reading test given at the end of the year if the Distar program successfully overcomes the lag in readiness.

Second Grade Testing

A non-Distar group was established at E.B. NEWTON in 1972-73 and compared to the Distar group at Newton School in that evaluation. Selection of children for the Distar and non-Distar classes was done in a way that would form two similar groups. The manner of selection of children for Distar instruction at Chase School was not expected to produce a comparable group. Although the three groups are compared in Tables A-1 and A-4, only comparisons of the Newton School Distar and control second grade classes can yield conclusions about the relative effectiveness of the two programs at second grade level.

It was stated in the 1973 report (p. 10) that reading test scores of the Chase School Distar group seemed extraordinarily high and the question of validity was raised. Reading scores of these children (Table A-4) are examined to resolve the question.

Table A-1 shows the two groups at Newton School are equivalent in IQ. However, seven children in the control group were tested in reading, but no IQ score was reported. IQ was reported on one child who was absent for the reading test. Five children who had been part of the Distar group at Newton in 1972-73 were no longer in the group in 1973-74. Table A-4 shows the two classes at Newton School did not differ significantly on any reading subtest score.

The question of exceptionally high performance by Chase School pupils in 1972-73 was reexamined. A comparison of achievement by this group and the other two shows no superior results by the Chase group in the interim testing (Table A-4). Of the original group in 1972-73, one child was not tested in the interim testing. Two new children, not present in 1972-73, were tested. After excluding these two, mean scores of the seventeen original children tested for this interim evaluation do not differ appreciably from the mean scores

of all nineteen reported in Table A-4. The Chase School group includes three children who had been transfered to traditional programs at the beginning of the year because of high performance. These children were given typewritten copies of the reading test, as was the control group at Newton. All children still receiving instruction with the Distar materials were given hand-lettered versions using the Distar alphabet form. Table A-4 shows that the Chase School group was in no way superior to the other groups in reading performance.

Self-Concept Measure

A measure of attitude about self entitled When Do I Smile? was used in November 1973. Groups were compared on this test and will again be compared in May. In addition, improvement in self-concept within each group will be assessed in May. Comparison of groups on the pretest shows only one significant difference. The comparison of first-grade groups at Chase and Dalrymple favors the latter (Table A-5; lower scores indicate higher self-concept). Unfortunately, almost half the response sheets of one class were invalidated by multiple responses or omitted responses.

Parent Questionnaire

Parents of Distar children and control groups (selected classes at Newton School) were asked to respond to a questionnaire. Responses of parents of boys and of girls at each school and grade were rather similarly distributed among the choices, therefore responses of boys and girls are combined - no differentiation is made by sex in Table A-6. In most cases questionnaires were completed by mothers. Of the ninety questionnaires returned, only five were completed by fathers or by both parents. Therefore, no differentiation is made as to which parent responded. Question 1 asked infor-

mation about child's grade, sex, and which parent responded.

Question 2 inquired about interest in reading. All groups reported that most children showed some interest or high interest.

Question 3 inquired about interest in arithmetic. All groups reported that most children showed some interest or high interest.

Question 4 inquired about the child's feelings about school. All groups generally reported that children enjoyed school.

Asked whether the parent was satisfied with the child's progress in reading (question 5), most responses in five groups were affirmative. Most parents of children in the first grade Distar class in one school responded that they were uncertain. In many of these cases, parents' commented that too little time in the program prevented them from making a judgment.

Missing data prevented analysis of responses by parents of second graders to questions 6, 7, and 8.

Asked about satisfaction with progress in arithmetic (question 6), responses were mainly affirmative in two groups and uncertain in one. Again, the response of uncertain was attributed to lack of sufficient time to make a judgment.

In most cases parents were satisfied with the teacher's response to the child and the child's relations with other children.(questions 7 and 8).

In conclusion, there was no group of children whose parents had primarily negative responses to questions about children's interest, progress, and relations to school, teacher, and peers. Neither the Distar nor the traditional programs at either first or second grade produced a fundamentally negative assessment by the parents to any of these questions.

Conclusions to Interim Report

1. Comparison of the first-grade classes at Chase, Newton and Dalrymple schools in IQ, oral language, and readiness shows the generally superior performance of the children at Dalrymple. This underscores the need for compensatory services for the children at the two Title I target schools.

2. Reading test results of three second-grade classes at Title I target schools fail to establish any superiority of Distar or traditional instruction over the other. There is evidence from classroom observation that second grade instruction of Distar classes increasingly consists of an amalgamation of Distar and traditional materials and procedures. Effectiveness of instruction during these first two years may result largely from the use of instructional aides. If this is so, equally good results can possibly be obtained by using an instructional program more flexible than Distar while maintaining the use of instructional aides and other support services.

3. Attitude measures and parent questionnaires tend to show equivalent attitudes about self, school, and the academic skills of reading and arithmetic.

TABLE A-1

COMPARISON OF IQ SCORES
JANUARY 1974

Group	N	Mean	SD	t	P
Grade 1					
Newton-Distar	21	101.7	17.70		
vs.				-0.145	NS
Chase-Distar	21	102.4	11.33		
Newton-Distar	21	101.7	17.70		
vs.				-1.858	NS
Dalrym.-Contr.	26	109.2	9.58		
Chase-Distar	21	102.4	11.33		
vs.				-2.246	<.05
Dalrym.-Contr.	26	109.2	9.58		
Grade 2					
Newton-Distar	23	108.5	15.08		
vs.				2.412	<.05
Chase-Distar	17	98.0	11.20		
Newton-Distar	23	108.5	15.08		
vs.				-0.395	NS
Newton-Control	20	110.2	13.26		
Chase-Distar	17	98.0	11.20		
vs.				-2.992	<.01
Newton-Control	20	110.2	13.26		

TABLE A-2

COMPARISON OF METROPOLITAN READINESS TEST
SCORES OF FIRST-GRADERS
(SEPTEMBER 1973)

Group	N	Mean	SD	t	P
Newton-Distar	20	50.40	15.55		
vs.				-0.873	NS
Chase-Distar	22	54.05	11.36		
Newton-Distar	20	50.40	15.55		
vs.				-3.660	<.001
Dalrym.-Contr.	25	64.56	10.32		
Chase-Distar	22	54.05	11.36		
vs.				-3.325	<.01
Dalrym.-Contr.	25	64.56	10.32		

TABLE A-3

COMPARISON OF ORAL LANGUAGE TEST SCORES OF FIRST-GRADERS (OCTOBER 1973)

Group	N	Mean	SD	t	P
Newton-Distar	21	7.86	2.61		
vs.				-0.654	NS
Chase-Distar	20	8.35	2.18		
Newton-Distar	21	7.86	2.61		
vs.				-2.970	<.01
Dalrym.-Contr.	22	10.18	2.52		
Chase-Distar	20	8.35	2.18		
vs.				-2.506	<.05
Dalrym.-Contr.	22	10.18	2.52		

TABLE A-4

COMPARISON OF READING SCORES ON NEW DISTAR READING TEST (SECOND-GRADERS, DECEMBER 1973)

Score	Newton-Distar (N = 23) Mean & SD	Chase-Distar (N = 19) Mean & SD	t	P
Main Idea	4.48 1.53	4.11 1.37	0.823	NS
Details	10.83 1.27	9.84 1.71	2.142	<.05
Inferences	4.57 .90	4.21 1.18	1.106	NS
Total Reading	19.87 3.28	18.16 3.39	1.659	NS

Score	Newton-Distar (N = 23) Mean & SD	Newton-Control (N = 26) Mean & SD	t	P
Main Idea	4.48 1.53	4.00 1.55	1.084	NS
Details	10.83 1.27	10.08 2.40	1.341	NS
Inferences	4.57 .90	4.15 1.26	1.305	NS
Total Reading	19.87 3.28	18.23 4.27	1.491	NS

TABLE A-4 (Continued)

Score	Chase-Distar (N = 19) Mean & SD	Newton-Control (N = 26). Mean & SD	t	P
Main Idea	4.11 1.37	4.00 1.55	0.236	NS
Details	9.84 1.71	10.08 2.40	-0.364	NS
Inferences	4.21 1.18	4.15 1.26	0.153	NS
Total Reading	18.16 3.39	18.23 4.27	-0.062	NS

TABLE A-5

COMPARISON OF GROUPS ON SELF-CONCEPT MEASURE WHEN DO I SMILE? (NOVEMBER 1973)

Group	N	Mean	SD	t	P
Grade 1					
Newton-Distar	12	38.08	16.60	-0.947	NS
vs.					
Chase-Distar	21	42.71	11.46		
Newton-Distar	12	38.08	16.60	1.318	NS
vs.					
Dalrym.-Contr.	23	32.35	9.29		
Chase-Distar	21	42.71	11.46	3.308	<.01
vs.					
Dalrym.-Contr.	23	32.35	9.29		
Grade 2					
Newton-Distar	22	39.06	12.61	-0.211	NS
vs.					
Chase-Distar	18	39.78	8.33		
Newton-Distar	22	39.06	12.61	0.259	NS
vs.					
Newton-Control	25	38.20	9.76		
Chase-Distar	18	39.78	8.33	0.555	NS
vs.					
Newton-Control	25	38.20	9.76		

TABLE A-6

TALLIES OF RESPONSES TO PARENT QUESTIONNAIRE

Question	Grade 1			Grade 2		
	Newton Distar	Chase Distar	Newton Control	Newton Distar	Chase Distar	Newton Control
2. interest in rdg.						
High	5	10	9	9	5	13
Some	6	3	7	9	5	5
Low	4	0	0	0	0	0
3. interest in arith.						
High	6	4	12	8	5	11
Some	4	9	4	9	5	6
Low	2	0	0	1	0	1
4. attitude to. school						
Enjoys	10	12	14	16	9	17
Neutral	2	2	2	1	1	1
Dislikes	1	0	0	1	0	0
5. satisfac. with rdg. progress						
Yes	3	12	12	16	8	13
Uncert.	9	2	3	2	2	5
No	1	0	1	0	0	0
6. Satisfac. with arith. progress						
Yes	4	10	16			
Uncert.	7	4	0			
No	1	0	0			
7. Satisfac. with teacher behav. to child.						
Yes	12	14	16			
Uncert.	2	0	0			
No	0	0	0			
8. Satisfac. with child's relations to other children						
Yes	9	10	16			
Uncert.	4	4	0			
No	1	0	0			
No. of boys	8	8	5	4	4	9
No. of girls	6	6	11	14	6	9

APPENDIX B

Readministration of Reading Test to Former First-Graders in October 1974

06040

In October 1974, second-graders (first-graders in 1973-1974) were readministered the New Distar Reading Test in order to re-check the ambiguous status of first-grade reading results. This step was in keeping with the first recommendation made on page 24 of the 1973-1974 report. It was thought that results obtained on the end-of-first-grade testing may have been caused by non-standard procedures of administering the tests to the three first-grade classes. For this reason, one reading specialist was assigned to administer the test to the three second-grade classes, with the assistance of the remedial reading teacher of the respective school.

The results of this testing are shown in Table B-1. The number of children tested in October 1974 in each class is not identical to the number tested in May 1974. In October, no significant differences appeared among the classes tested, whereas significant differences had occurred in May 1974 (Table 3, 1973-1974 report). When data were analyzed on only those children tested both in May and in October, the significant differences that occurred for these children on the May testing were on the same comparisons on which significant differences appeared for the total number tested in May, whereas in October, only one significant difference was obtained (Tables B-2 and B-3). A comparison of the May and October results within each group (Table B-4) shows significant gains made by the Newton-Distar group on all scores, significant gains made by the Dalrymple-control group on two scores, and significant decline made by the Chase-Distar group on two scores. These results suggest that non-standard testing procedures were partly responsible for the May results. When procedures were standardized, and after several months passed in which the children matured and received some additional instruction, the groups did not differ appreciably in reading ability. In conclusion, the Distar reading program was

judged to be neither better nor worse than conventional instruction given to first-graders.

TABLE B-1

COMPARISON OF SCORES ON OCTOBER 1974 ADMINISTRATION OF NEW DISTAR READING TEST, ALL CHILDREN TESTED

Score	Newton-Distar (N = 19) Mean & SD	Chase-Distar (N = 21) Mean & SD	t	P*
Main Idea	2.37 2.11	3.43 1.43	-1.872	NS
Details	6.32 4.69	7.00 3.00	-0.555	NS
Inferences	2.68 2.19	2.90 1.30	-0.392	NS
Total Reading	11.37 8.47	13.33 4.37	-0.935	NS

Score	Newton-Distar (N = 19) Mean & SD	Dalrymple-Control (N = 24) Mean & SD	t	P*
Main Idea	2.37 2.11	3.46 1.56	-1.946	NS
Details	6.32 4.69	8.21 3.58	-1.502	NS
Inferences	2.68 2.19	3.42 1.53	-1.291	NS
Total Reading	11.37 8.47	15.08 6.02	-1.680	NS

Score	Chase-Distar (N = 21) Mean & SD	Dalrymple-Control (N = 24) Mean & SD		
Main Idea	3.43 1.43	3.46 1.56	-0.066	NS
Details	7.00 3.00	8.21 3.58	-1.218	NS
Inferences	2.90 1.30	3.42 1.53	-1.200	NS
Total Reading	13.33 4.37	15.08 6.02	-1.101	NS

*NS = Not Significant

TABLE B-2

COMPARISON OF SCORES ON MAY 1974 ADMINISTRATION OF NEW DISTAR READING TEST FOR CHILDREN TESTED IN BOTH MAY AND OCTOBER

Score	Newton-Distar (N = 19) Mean & SD	Chase-Distar (N = 19) Mean & SD	t	P*
Main Idea	1.32 1.80	4.32 1.11	-6.195	<.001
Details	1.84 2.61	10.16 1.42	-12.194	<.001
Inferences	.95 1.39	3.68 1.25	-6.374	<.001
Total Reading	4.11 4.64	18.16 2.41	-11.713	<.001

Score	Newton-Distar (N = 19) Mean & SD	Dalrymple-Control (N = 22) Mean & SD	t	P*
Main Idea	1.32 1.80	2.95 1.59	-3.101	<.01
Details	1.84 2.61	6.55 3.08	-5.227	<.001
Inferences	.95 1.39	3.14 1.46	-4.894	<.001
Total Reading	4.11 4.64	12.64 5.08	-5.582	<.001

Score	Chase-Distar (N = 19) Mean & SD	Dalrymple-Control (N = 22) Mean & SD	t	P*
Main Idea	4.32 1.11	2.95 1.59	3.133	<.01
Details	10.16 1.42	6.55 3.08	4.689	<.001
Inferences	3.68 1.25	3.14 1.46	1.281	NS
Total Reading	18.16 2.41	12.64 5.08	4.333	<.001

NS*= Not Significant

TABLE B-3

COMPARISON OF SCORES ON OCTOBER 1974 ADMINISTRATION OF NEW DISTAR READING TEST FOR CHILDREN TESTED IN BOTH MAY AND OCTOBER

Score	Newton-Distar (N = 19) Mean & SD	Chase-Distar (N = 19) Mean & SD	t	P*
Main Idea	2.37 2.11	3.63 1.26	-2.239	<.05
Details	6.32 4.69	7.53 2.61	-0.983	NS
Inferences	2.68 2.19	2.95 1.35	-0.446	NS
Total Reading	11.37 8.47	14.11 3.70	-1.291	NS

Score	Newton-Distar (N = 19) Mean & SD	Dalrymple-Control (N = 22) Mean & SD	t	P*
Main Idea	2.37 2.11	3.36 1.53	-1.744	NS
Details	6.32 4.69	7.91 3.58	-1.231	NS
Inferences	2.68 2.19	3.32 1.55	-1.081	NS
Total Reading	11.37 8.47	14.59 6.05	-1.415	NS

Score	Chase-Distar (N = 19) Mean & SD	Dalrymple-Control (N = 22) Mean & SD	t	P*
Main Idea	3.63 1.26	3.36 1.53	0.607	NS
Details	7.53 2.61	7.91 3.58	-0.385	NS
Inferences	2.95 1.35	3.32 1.55	-0.808	NS
Total Reading	14.11 3.70	14.59 6.05	-0.304	NS

*NS = Not Significant

TABLE B-4

COMPARISON OF MAY 1974 AND OCTOBER 1974
ADMINISTRATIONS OF NEW DISTAR READING TEST

Score	Pretest Mean & SD	Posttest Mean & SD	t	P*
	Newton-Distar, N = 19			
Main Idea	1.32 1.80	2.37 2.11	2.344	<.05
Details	1.84 2.61	6.32 4.69	4.819	<.001
Inferences	.95 1.39	2.68 2.19	3.598	<.01
Total Reading	4.11 4.64	11.37 8.47	4.692	<.001
	Chase-Distar, N = 19			
Main Idea	4.32 1.11	3.63 1.26	-1.660	NS
Details	10.16 1.42	7.53 2.61	-3.773	<.01
Inferences	3.68 1.25	2.95 1.35	-1.973	NS
Total Reading	18.16 2.41	14.11 3.70	-4.095	<.001
	Dalrymple-Control, N = 22			
Main Idea	2.95 1.59	3.36 1.53	1.056	NS
Details	6.55 3.08	7.91 3.58	2.485	<.05
Inferences	3.14 1.46	3.32 1.55	0.748	NS
Total Reading	12.64 5.08	14.59 6.05	2.506	<.05

*NS = Not Significant